

Quantitative estimates in approximation by Bernstein-Durrmeyer-Choquet operators with respect to monotone and submodular set functions

Sorin G. Gal

Department of Mathematics and Computer Science,
University of Oradea,
Universitatii Street No.1, 410087, Oradea, Romania
E-mail: *galso@uoradea.ro*

Abstract

For the qualitative results of pointwise and uniform approximation obtained in [10], we present general quantitative estimates in terms of the modulus of continuity and in terms of a K -functional, for the generalized multivariate Bernstein-Durrmeyer operator $M_{n,\Gamma_{n,x}}$, written in terms of the Choquet integral with respect to a family of monotone and submodular set functions, $\Gamma_{n,x}$, on the standard d -dimensional simplex. When $\Gamma_{n,x}$ reduces to two elements, one a Choquet submodular set function and the other one a Borel measure, for suitable modified Bernstein-Durrmeyer operators, univariate L^p -approximations, $p \geq 1$, with estimates in terms of a K -functional are proved. In the particular cases when $d = 1$ and the Choquet integral is taken with respect to some concrete possibility measures, the pointwise estimate in terms of the modulus of continuity is detailed. Some simple concrete examples of operators improving the classical error estimates are presented. Potential applications to practical methods dealing with data, like learning theory and regression models, also are mentioned.

AMS 2000 Mathematics Subject Classification: 41A36, 41A25, 28A12, 28A25.

Keywords and phrases: Bernstein-Durrmeyer-Choquet operator, possibility measure, monotone and submodular set function, Choquet integral, quantitative estimates, pointwise approximation, uniform approximation, L^p -approximation, modulus of continuity, K -functional.

1 Introduction

The approximation properties of the multivariate Bernstein-Durrmeyer linear operator defined with respect to a nonnegative, bounded Borel measure $\mu : \mathcal{B}_{S^d} \rightarrow \mathbb{R}_+$, by

$$M_{n,\mu}(f)(x) = \sum_{|\alpha|=n} \frac{\int_{S^d} f(t) B_\alpha(t) d\mu(t)}{\int_{S^d} B_\alpha(t) d\mu(t)} \cdot B_\alpha(x) := \sum_{|\alpha|=n} c(\alpha, \mu) \cdot B_\alpha(x), \quad x \in S^d, \quad n \in \mathbb{N}, \quad (1)$$

where \mathcal{B}_{S^d} denotes the sigma algebra of all Borel measurable subsets in the power set $\mathcal{P}(S^d)$ and f is supposed to be μ -integrable on the standard simplex

$$S^d = \{(x_1, \dots, x_d); 0 \leq x_1, \dots, x_d \leq 1, 0 \leq x_1 + \dots + x_d \leq 1\},$$

were studied in, e.g., the recent papers [4], [1], [2], [3] and [13].

Note that in (1), it is used the notation

$$B_\alpha(x) = \frac{n!}{\alpha_0! \cdot \alpha_1! \cdot \dots \cdot \alpha_d!} (1 - x_1 - x_2 - \dots - x_d)^{\alpha_0} \cdot x_1^{\alpha_1} \cdot \dots \cdot x_d^{\alpha_d} \\ =: \frac{n!}{\alpha_0! \cdot \alpha_1! \cdot \dots \cdot \alpha_d!} \cdot P_\alpha(x),$$

where $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_d)$, $\alpha_j \in \mathbb{N} \cup \{0\}$, $j = 0, \dots, d$, $|\alpha| = \alpha_0 + \alpha_1 + \dots + \alpha_d = n$.

In the very recent paper [10], we have proved that the approximation results in the above mentioned papers remain valid for the more general case when μ is a monotone, normalized and submodular set function on S^d and the integrals used in (1) are the nonlinear Choquet integrals with respect to μ .

The main goal of this paper is to obtain quantitative estimates in terms of the modulus of continuity and in terms of some K -functionals, for the pointwise and uniform approximation obtained in [10] and for the univariate L^p -approximation, $p \geq 1$, in the case of more general multivariate Bernstein-Durrmeyer polynomial operators defined by

$$M_{n,\Gamma_{n,x}}(f)(x) = \sum_{|\alpha|=n} c(\alpha, \mu_{n,\alpha,x}) \cdot B_\alpha(x), \quad x \in S^d, \quad n \in \mathbb{N}, \quad (2)$$

where

$$c(\alpha, \mu_{n,\alpha,x}) = \frac{(C) \int_{S^d} f(t) B_\alpha(t) d\mu_{n,\alpha,x}(t)}{(C) \int_{S^d} B_\alpha(t) d\mu_{n,\alpha,x}(t)} = \frac{(C) \int_{S^d} f(t) P_\alpha(t) d\mu_{n,\alpha,x}(t)}{(C) \int_{S^d} P_\alpha(t) d\mu_{n,\alpha,x}(t)}$$

and for every $n \in \mathbb{N}$ and $x \in S^d$, $\Gamma_{n,x} = (\mu_{n,\alpha,x})_{|\alpha|=n}$ is a family of bounded, monotone, submodular and strictly positive set functions on \mathcal{B}_{S^d} .

Note that if $\Gamma_{n,x}$ reduces to one element (i.e. $\mu_{n,\alpha,x} = \mu$ for all n , x and α), then the operator given by (2) reduces to the operator considered in [10].

The plan of the paper is as follows. Section 2 contains some preliminaries on possibility theory and on Choquet integral. In Section 3, general quantitative estimates in terms of the modulus of continuity and in terms of a

K -functional for the pointwise and uniform approximation by the operators $M_{n,\Gamma_{n,x}}(f)(x)$ defined by (2) are obtained. Also, when $\Gamma_{n,x}$ reduces to two elements, one a Choquet submodular set function and the other one a Borel measure, for suitable modified Bernstein-Durrmeyer-Choquet operators, univariate L^p -approximations, $p \geq 1$, with quantitative estimates in terms of a K -functional are presented. Finally, in Section 4, in the particular case when $d = 1$ and the Choquet integrals are taken with respect to some concrete possibility measures, the pointwise estimate in terms of the modulus of continuity is detailed. Also, some concrete example of operators improving the classical error estimates are presented and potential applications to practical methods dealing with data are mentioned.

2 Preliminaries

Firstly, we present a few known concepts in possibility theory useful for the next considerations. For details, see, e.g., [8].

Definition 2.1. For the non-empty set Ω , denote by $\mathcal{P}(\Omega)$ the family of all subsets of Ω .

(i) A function $\lambda : \Omega \rightarrow [0, 1]$ with the property $\sup\{\lambda(s); s \in \Omega\} = 1$, is called possibility distribution on Ω .

(ii) A possibility measure is a set function $P : \mathcal{P}(\Omega) \rightarrow [0, 1]$, satisfying the axioms $P(\emptyset) = 0$, $P(\Omega) = 1$ and $P(\bigcup_{i \in I} A_i) = \sup\{P(A_i); i \in I\}$ for all $A_i \subset \Omega$, and any I , an at most countable family of indices. Note that if $A, B \subset \Omega$, $A \subset B$, then the last property easily implies that $P(A) \leq P(B)$ and that $P(A \cup B) \leq P(A) + P(B)$.

Any possibility distribution λ on Ω , induces the possibility measure $P_\lambda : \mathcal{P}(\Omega) \rightarrow [0, 1]$, given by the formula $P_\lambda(A) = \sup\{\lambda(s); s \in A\}$, for all $A \subset \Omega$ (see, e.g., [8], Chapter 1).

Some known concepts and results concerning the Choquet integral can be summarized by the following.

Definition 2.2. Suppose $\Omega \neq \emptyset$ and let \mathcal{C} be a σ -algebra of subsets in Ω .

(i) (see, e.g., [17], p. 63) The set function $\mu : \mathcal{C} \rightarrow [0, +\infty]$ is called a monotone set function (or capacity) if $\mu(\emptyset) = 0$ and $\mu(A) \leq \mu(B)$ for all $A, B \in \mathcal{C}$, with $A \subset B$. Also, μ is called submodular if

$$\mu(A \cup B) + \mu(A \cap B) \leq \mu(A) + \mu(B), \text{ for all } A, B \in \mathcal{C}.$$

μ is called bounded if $\mu(\Omega) < +\infty$ and normalized if $\mu(\Omega) = 1$.

(ii) (see, e.g., [17], p. 233, or [6]) If μ is a monotone set function on \mathcal{C} and if $f : \Omega \rightarrow \mathbb{R}$ is \mathcal{C} -measurable (that is, for any Borel subset $B \subset \mathbb{R}$ it follows $f^{-1}(B) \in \mathcal{C}$), then for any $A \in \mathcal{C}$, the concept of Choquet integral is defined by

$$(C) \int_A f d\mu = \int_0^{+\infty} \mu(F_\beta(f) \cap A) d\beta + \int_{-\infty}^0 [\mu(F_\beta(f) \cap A) - \mu(A)] d\beta,$$

where we used the notation $F_\beta(f) = \{\omega \in \Omega; f(\omega) \geq \beta\}$. Notice that if $f \geq 0$ on A , then in the above formula we get $\int_{-\infty}^0 = 0$.

The function f will be called Choquet integrable on A if $(C) \int_A f d\mu \in \mathbb{R}$.

In what follows, we list some known properties of the Choquet integral.

Remark 2.3. If $\mu : \mathcal{C} \rightarrow [0, +\infty]$ is a monotone set function, then the following properties hold :

(i) For all $a \geq 0$ we have $(C) \int_A a f d\mu = a \cdot (C) \int_A f d\mu$ (if $f \geq 0$ then see, e.g., [17], Theorem 11.2, (5), p. 228 and if f is of arbitrary sign, then see, e.g., [7], p. 64, Proposition 5.1, (ii)).

(ii) For all $c \in \mathbb{R}$ and f of arbitrary sign, we have (see, e.g., [17], pp. 232-233, or [7], p. 65) $(C) \int_A (f + c) d\mu = (C) \int_A f d\mu + c \cdot \mu(A)$.

If μ is submodular too, then for all f, g of arbitrary sign and lower bounded, we have (see, e.g., [7], p. 75, Theorem 6.3)

$$(C) \int_A (f + g) d\mu \leq (C) \int_A f d\mu + (C) \int_A g d\mu.$$

(iii) If $f \leq g$ on A then $(C) \int_A f d\mu \leq (C) \int_A g d\mu$ (see, e.g., [17], p. 228, Theorem 11.2, (3) if $f, g \geq 0$ and p. 232 if f, g are of arbitrary sign).

(iv) Let $f \geq 0$. If $A \subset B$ then $(C) \int_A f d\mu \leq (C) \int_B f d\mu$. In addition, if μ is finitely subadditive, then

$$(C) \int_{A \cup B} f d\mu \leq (C) \int_A f d\mu + (C) \int_B f d\mu.$$

(v) It is immediate that $(C) \int_A 1 \cdot d\mu(t) = \mu(A)$.

(vi) The formula $\mu(A) = \gamma(M(A))$, where $\gamma : [0, 1] \rightarrow [0, 1]$ is an increasing and concave function, with $\gamma(0) = 0$, $\gamma(1) = 1$ and M is a probability measure (or only finitely additive) on a σ -algebra on Ω (that is, $M(\emptyset) = 0$, $M(\Omega) = 1$ and M is countably additive), gives simple examples of normalized, monotone and submodular set functions (see, e.g., [7], pp. 16-17, Example 2.1). For example, we can take $\gamma(t) = \sqrt{t}$.

If the above γ function is increasing, concave and satisfies only $\gamma(0) = 0$, then for any bounded Borel measure m , $\mu(A) = \gamma(m(A))$ gives a simple example of bounded, monotone and submodular set function.

Note that any possibility measure μ is normalized, monotone and submodular. Indeed, the axiom $\mu(A \cup B) = \max\{\mu(A), \mu(B)\}$ implies the monotonicity, while the property $\mu(A \cap B) \leq \min\{\mu(A), \mu(B)\}$ implies the submodularity.

(vii) If μ is a countably additive bounded measure, then the Choquet integral $(C) \int_A f d\mu$ reduces to the usual Lebesgue type integral (see, e.g., [7], p. 62, or [17], p. 226).

3 Quantitative estimates for general Bernstein-Durrmeyer-Choquet operators

Recall that $\mu : \mathcal{B}_{S^d} \rightarrow [0, +\infty)$ is said strictly positive if for every open set $A \subset \mathbb{R}^n$ with $A \cap S^d \neq \emptyset$, we have $\mu(A \cap S^d) > 0$.

The support of μ is defined by

$$\text{supp}(\mu) = \{x \in S^d; \mu(N_x) > 0 \text{ for every open neighborhood } N_x \in \mathcal{B}_{S^d} \text{ of } x\}.$$

Note that the strict positivity of μ , evidently implies the condition $\text{supp}(\mu) \setminus \partial S^d \neq \emptyset$, which guarantees that $(C) \int_{S^d} B_\alpha(t) d\mu(t) > 0$, for all B_α .

Let us consider $C_+(S^d) = \{f : S^d \rightarrow \mathbb{R}_+; f \text{ is continuous on } S^d\}$, endowed with the norm $\|F\|_{C(S^d)} = \sup\{|F(x)|; x \in S^d\}$.

The first main result of this section consists in the following general quantitative estimates in pointwise and uniform approximation.

Theorem 3.1. *For each fixed $n \in \mathbb{N}$ and $x \in S^d$, let $\Gamma_{n,x} = \{\mu_{n,\alpha,x}\}_{|\alpha|=n}$ be a family of bounded, monotone, submodular and strictly positive set functions on \mathcal{B}_{S^d} .*

(i) *For every $f \in C_+(S^d)$, $x = (x_1, \dots, x_d) \in S^d$, $n \in \mathbb{N}$, we have*

$$|M_{n,\Gamma_{n,x}}(f)(x) - f(x)| \leq 2\omega_1(f; M_{n,\Gamma_{n,x}}(\varphi_x)(x))_{S^d},$$

where $M_{n,\Gamma_{n,x}}(f)(x)$ is given by (2), $\|x\| = \sqrt{\sum_{i=1}^d x_i^2}$, $\varphi_x(t) = \|t - x\|$ and $\omega_1(f; \delta)_{S^d} = \sup\{|f(t) - f(x)|; t, x \in S^d, \|t - x\| \leq \delta\}$.

(ii) *Suppose that the family $\Gamma_{n,x}$ does not depend on x . Then, for any $f \in C_+(S^d)$ and $n \in \mathbb{N}$, we get*

$$\|M_{n,\Gamma_n}(f) - f\|_{C(S^d)} \leq 2K\left(f; \frac{\Delta_n}{2}\right),$$

where $\Delta_n = \sum_{i=1}^d \|M_{n,\Gamma_n}(|\varphi_{e_i} - x_i \mathbf{1}|)\|_{C(S^d)}$,

$$K(f; t) = \inf_{g \in C_+^1(S^d)} \{\|f - g\|_{C(S^d)} + t\|\nabla g\|_{C(S^d)}\},$$

$C_+^1(S^d)$ is the subspace of all functions $g \in C_+(S^d)$ with continuous partial derivatives $\partial_i g$, $i \in \{1, \dots, d\}$ and

$$\|\nabla g\|_{C(S^d)} = \max_{i \in \{1, \dots, d\}} \{\|\partial_i g\|_{C(S^d)}\},$$

$\varphi_{e_i}(x) = x_i$, $i \in \{1, \dots, d\}$, $x = (x_1, \dots, x_d)$, $\mathbf{1}(x) = 1$, for all $x \in S^d$.

Proof. (i) For $x \in S^d$, $n \in \mathbb{N}$ and $|\alpha| = n$ arbitrary fixed, let us consider $T_{n,\alpha,x} : C_+(S^d) \rightarrow \mathbb{R}_+$ defined by

$$T_{n,\alpha,x}(f) = (C) \int_{S^d} f(t) P_\alpha(t) d\mu_{n,\alpha,x}(t), f \in C_+(S^d).$$

Based on the above Remark 2.3, (i), (ii), (iii) and reasoning exactly as in the proof of Lemma 3.1 in [10], we get $|T_{n,\alpha,x}(f) - T_{n,\alpha,x}(g)| \leq T_{n,\alpha,x}(|f - g|)$. Then, since $T_{n,\alpha,x}$ is positively homogeneous, sublinear and monotonically increasing, we immediately get that $M_{n,\Gamma_{n,x}}$ keeps the same properties and as a consequence it follows

$$|M_{n,\Gamma_{n,x}}(f)(x) - M_{n,\Gamma_{n,x}}(g)(x)| \leq M_{n,\Gamma_{n,x}}(|f - g|)(x), \quad (3)$$

$M_{n,\Gamma_{n,x}}(\lambda f) = \lambda M_{n,\Gamma_{n,x}}(f)$, $M_{n,\Gamma_{n,x}}(f+g) \leq M_{n,\Gamma_{n,x}}(f) + M_{n,\Gamma_{n,x}}(g)$ and that $f \leq g$ on S^d implies $M_{n,\Gamma_{n,x}}(f) \leq M_{n,\Gamma_{n,x}}(g)$ on S^d , for all $\lambda \geq 0$, $f, g \in C_+(S^d)$, $n \in \mathbb{N}$, $|\alpha| = n$, $x \in S^d$.

Denoting $e_0(t) = 1$ for all $t \in S^d$, since obviously $M_{n,\Gamma_{n,x}}(e_0)(x) = 1$ for all $x \in S^d$ and taking into account the properties in Remark 2.3, (i) and (3), for any fixed x we obtain

$$\begin{aligned} |M_{n,\Gamma_{n,x}}(f)(x) - f(x)| &= |M_{n,\Gamma_{n,x}}(f(t))(x) - M_{n,\Gamma_{n,x}}(f(x))(x)| \\ &\leq M_{n,\Gamma_{n,x}}(|f(t) - f(x)|)(x). \end{aligned} \quad (4)$$

But taking into account the properties of the modulus of continuity, for all $t, x \in S^d$ and $\delta > 0$, we get

$$|f(t) - f(x)| \leq \omega_1(f; \|t - x\|)_{S^d} \leq \left[\frac{1}{\delta} \|t - x\| + 1 \right] \omega_1(f; \delta)_{S^d}. \quad (5)$$

Now, from (4) and applying $M_{n,\Gamma_{n,x}}$ to (5), by the properties of $M_{n,\Gamma_{n,x}}$ mentioned after the inequality (3), we immediately get

$$|M_{n,\Gamma_{n,x}}(f)(x) - f(x)| \leq \left[\frac{1}{\delta} M_{n,\Gamma_{n,x}}(\varphi_x)(x) + 1 \right] \omega_1(f; \delta)_{S^d}.$$

Choosing here $\delta = M_{n,\Gamma_{n,x}}(\varphi_x)(x)$, we obtain the desired estimate.

(ii) Let $f, g \in C_+(S^d)$. We have

$$\begin{aligned} &f(x) - M_{n,\Gamma_n}(f)(x) \\ &= f(x) - g(x) + M_{n,\Gamma_n}(g)(x) - M_{n,\Gamma_n}(f)(x) + g(x) - M_{n,\Gamma_n}(g)(x), \end{aligned}$$

which, by using (3) too, implies

$$\begin{aligned} &|f(x) - M_{n,\Gamma_n}(f)(x)| \\ &\leq |f(x) - g(x)| + |M_{n,\Gamma_n}(g)(x) - M_{n,\Gamma_n}(f)(x)| + |g(x) - M_{n,\Gamma_n}(g)(x)| \\ &\leq |f(x) - g(x)| + M_{n,\Gamma_n}(|g - f|)(x) + |g(x) - M_{n,\Gamma_n}(g)(x)| \\ &\leq 2\|f - g\|_{C(S^d)} + |g(x) - M_{n,\Gamma_n}(g)(x)|. \end{aligned}$$

By following the lines in the proof of Theorem 4.5 in [4], since from the lines after relation (3) in the above point (i), the operator M_{n,Γ_n} is monotone and subadditive, for all $g \in C_+^1(S^d)$, $x \in S^d$, we immediately get

$$\begin{aligned} &|g(x) - M_{n,\Gamma_n}(g)(x)| \\ &\leq M_{n,\Gamma_n}(|g - g(x)\mathbf{1}|)(x) \leq \|\nabla g\|_{C(S^d)} \cdot M_{n,\Gamma_n} \left(\sum_{i=1}^d |\varphi_{e_i} - x_i \mathbf{1}| \right) (x) \\ &\leq \|\nabla g\|_{C(S^d)} \cdot \sum_{i=1}^d M_{n,\Gamma_n}(|\varphi_{e_i} - x_i \mathbf{1}|)(x) \leq \|\nabla g\|_{C(S^d)} \cdot \Delta_n. \end{aligned}$$

Concluding, it follows

$$\|f - M_{n,\Gamma_n}(f)\|_{C(S^d)} \leq 2 \left[\|f - g\|_{C(S^d)} + \frac{\Delta_n}{2} \|\nabla g\|_{C(S^d)} \right],$$

which immediately implies the required estimate in (ii). \square

Remark 3.2. The positivity of function f in Theorem 3.1, (i) and (ii), is necessary because of the positive homogeneity of the Choquet integral used in their proofs. However, if f is of arbitrary sign and lower bounded on S^d with $f(x) - m \geq 0$, for all $x \in S^d$, then the statement of Theorem 3.1, (i), (ii), can be restated for the slightly modified Bernstein-Durrmeyer operator defined by

$$M_{n,\Gamma_{n,x}}^*(f)(x) = M_{n,\Gamma_{n,x}}(f - m)(x) + m.$$

Indeed, in the case of Theorem 3.1, (i), this is immediate from $\omega_1(f - m; \delta)_{S^d} = \omega_1(f; \delta)_{S^d}$ and from $M_{n,\Gamma_{n,x}}^*(f)(x) - f(x) = M_{n,\Gamma_{n,x}}(f - m)(x) - (f(x) - m)$. Note that in the case of Theorem 3.1, (ii), since we may consider here that $m < 0$, we immediately get the relations

$$\begin{aligned} K(f - m; t) &= \inf_{g \in C_+^1(S^d)} \{ \|f - (g + m)\|_{C(S^d)} + t \|\nabla g\|_{C(S^d)} \} \\ &= \inf_{g \in C_+^1(S^d)} \{ \|f - (g + m)\|_{C(S^d)} + t \|\nabla(g + m)\|_{C(S^d)} \} \\ &= \inf_{h \in C^1(S^d), h \geq m} \{ \|f - h\|_{C(S^d)} + t \|\nabla h\|_{C(S^d)} \}. \end{aligned}$$

In the particular case when the family $\Gamma_{n,x}$ does not depend on x and n , it is natural to ask for quantitative estimates in the L^1 -approximation of the Choquet integrable functions (not necessarily continuous). If, for example, $\Gamma_{n,x} = \{\mu\}$, $d = 1$ and μ is a bounded, monotone and submodular set function, then for the Bernstein-Durrmeyer-Choquet operators

$$D_{n,\mu}(f)(x) = \sum_{k=0}^n p_{n,k}(x) \cdot \frac{(C) \int_0^1 f(t) p_{n,k}(t) d\mu(t)}{(C) \int_0^1 p_{n,k}(t) d\mu(t)}, \quad p_{n,k}(x) = \binom{n}{k} x^k (1-x)^{n-k},$$

with $f \in L_\mu^1$ meaning f is $\mathcal{B}_{[0,1]}$ measurable and $\|f\|_{L_\mu^1} = (C) \int_0^1 |f(t)| d\mu(t) < \infty$, we get

$$\|D_{n,\mu}(f)\|_{L_\mu^1} \leq \sum_{k=0}^n (C) \int_0^1 p_{n,k}(t) |f(t)| d\mu(t) \leq n \cdot \|f\|_{L_\mu^1}, \quad n \in \mathbb{N}.$$

This is due to the fact that $(C) \int f d\mu$ is not, in general, additive as function of f (it is only subadditive).

Therefore, quantitative estimates in L_μ^p -approximation by Bernstein-Durrmeyer-Choquet operators, remains, in the general case, an open question.

However, in the particular case when the family of set functions $\Gamma_{n,x}$ reduces, for example, to two elements (one being a Choquet submodular set function μ

and the other one a Borel measure δ), for suitable defined Bernstein-Durrmeyer-Choquet operators, quantitative L_μ^p -approximation results, $p \geq 1$ hold. For this purpose, let us make the following notations :

$$L_\mu^p = \{f : [0, 1] \rightarrow \mathbb{R}; f \text{ is } \mathcal{B}_{[0,1]} \text{ measurable and } (C) \int_0^1 |f(t)|^p d\mu(t) < +\infty\},$$

$$L_{\mu,+}^p = L_\mu^p \cap \{f : [0, 1] \rightarrow \mathbb{R}_+\},$$

$$\overline{K}(f; t)_{L_{\mu,\delta}^p} = \inf_{g \in C_+^1([0,1])} \{\|f - g\|_{L_\mu^p} + \|f - g\|_{L_\delta^p} + t\|g'\|_{C([0,1])}\}.$$

It is easy to see that if $\mu \leq \delta$ then for all $f \in L_\delta^1$ and $t \geq 0$, we have $2K(f; t/2)_{L_\mu^p} \leq \overline{K}(f; t)_{L_{\mu,\delta}^p} \leq 2K(f; t)_{L_\delta^p}$, where K is of usual form and with the infimum taken for $g \in C_+^1([0, 1])$.

For $p = 1$, we have :

Theorem 3.3. *Let μ be a bounded, monotone, submodular and strictly positive set function on $\mathcal{B}_{[0,1]}$ and δ a bounded, strictly positive Borel measure on $\mathcal{B}_{[0,1]}$, such that $\mu(A) \leq \delta(A)$ for all $A \in \mathcal{B}_{[0,1]}$. Then, denoting $L_{\delta,+}^1 \subset L_{\mu,+}^1$ and defining the Bernstein-Durrmeyer-Choquet operators*

$$\overline{D}_{n,\delta,\mu}(f)(x) = \sum_{k=0}^{n-1} p_{n,k}(x) \cdot \frac{\int_0^1 f(t) p_{n,k}(t) d\delta(t)}{\int_0^1 p_{n,k}(t) d\delta(t)} + x^n \cdot \frac{(C) \int_0^1 f(t) t^n d\mu(t)}{(C) \int_0^1 t^n d\mu(t)},$$

for all $f \in L_{\delta,+}^1$, $n \in \mathbb{N}$ and denoting $\varphi_x(t) = |t - x|$, we have

$$\|f - \overline{D}_{n,\delta,\mu}(f)\|_{L_\mu^1} \leq 2\overline{K}\left(f; \frac{\|\overline{D}_{n,\delta,\mu}(\varphi_x)\|_{L_\mu^1}}{2}\right)_{L_{\mu,\delta}^1}.$$

Proof. Firstly, note that δ is monotone, submodular (in fact modular, i.e. submodular with equality) strictly positive and that for all $f \in L_{\delta,+}^1$ we have $\int_0^1 f(t) d\delta(t) = (C) \int_0^1 f(t) d\delta(t)$ (see Remark 2.3, (vii)). From here, from $\mu \leq \delta$ and from Definition 2.2, (ii), we immediately get $(C) \int_0^1 f(t) d\mu(t) \leq \int_0^1 f(t) d\delta(t)$, which means $L_{\delta,+}^1 \subset L_{\mu,+}^1$ and for all $f \in L_{\delta,+}^1$ implies

$$\begin{aligned} \|\overline{D}_{n,\delta,\mu}(f)\|_{L_\mu^1} &\leq \sum_{k=0}^{n-1} \frac{(C) \int_0^1 p_{n,k}(x) d\mu(x)}{\int_0^1 p_{n,k}(t) d\delta(t)} \cdot \int_0^1 f(t) p_{n,k}(t) d\delta(t) + (C) \int_0^1 f(t) t^n d\mu(t) \\ &\leq \int_0^1 f(t) \left[\sum_{k=0}^{n-1} p_{n,k}(t) \right] d\delta(t) + (C) \int_0^1 f(t) t^n d\mu(t) \leq \|f\|_{L_\delta^1} + \|f\|_{L_\mu^1}. \end{aligned} \quad (6)$$

Let $f, g \in L_{\delta,+}^1$. From

$$\begin{aligned} &|f(t) - \overline{D}_{n,\delta,\mu}(f)(t)| \\ &\leq |f(t) - g(t)| + |\overline{D}_{n,\delta,\mu}(g)(t) - \overline{D}_{n,\delta,\mu}(f)(t)| + |g(t) - \overline{D}_{n,\delta,\mu}(g)(t)|, \end{aligned}$$

integrating with respect to μ , from the properties of the Choquet integral, of the operator $\overline{D}_{n,\delta,\mu}$ (similar with those of $M_{n,\Gamma_{n,x}}$ in the proof of Theorem 3.1, (i)) and from (6), we obtain

$$\begin{aligned}
\|f - \overline{D}_{n,\delta,\mu}\|_{L_\mu^1} &= (C) \int_0^1 |f(t) - \overline{D}_{n,\delta,\mu}(f)(t)| d\mu(t) \\
&\leq (C) \int_0^1 |f(t) - g(t)| d\mu(t) + (C) \int_0^1 |\overline{D}_{n,\delta,\mu}(g)(t) - \overline{D}_{n,\delta,\mu}(f)(t)| d\mu(t) \\
&\quad + (C) \int_0^1 |g(t) - \overline{D}_{n,\delta,\mu}(g)(t)| d\mu(t) \\
&\leq \|f - g\|_{L_\mu^1} + (C) \int_0^1 \overline{D}_{n,\delta,\mu}(|g - f|)(t) d\mu(t) + \|g - \overline{D}_{n,\delta,\mu}(g)\|_{L_\mu^1} \\
&\leq \|f - g\|_{L_\mu^1} + (\|f - g\|_{L_\mu^1} + \|f - g\|_{L_\delta^1}) + \|g - \overline{D}_{n,\delta,\mu}(g)\|_{L_\mu^1}.
\end{aligned}$$

It remains to estimate $\|g - \overline{D}_{n,\delta,\mu}(g)\|_{L_\mu^1}$. But from

$$|g(x) - \overline{D}_{n,\delta,\mu}(g)(x)| = |\overline{D}_{n,\delta,\mu}(g(x))(x) - \overline{D}_{n,\delta,\mu}(g(t))(x)| \leq \overline{D}_{n,\delta,\mu}(|g(x) - g(t)|)(x)$$

and since for $g \in C_+^1([0, 1])$, we get $|g(x) - g(t)| \leq \|g'\|_{C([0,1])}|x - t|$, applying $\overline{D}_{n,\delta,\mu}$, it follows $\overline{D}_{n,\delta,\mu}(|g(x) - g(t)|)(x) \leq \|g'\|_{C([0,1])} \cdot \overline{D}_{n,\delta,\mu}(\varphi_x)(x)$.

Therefore, integrating above with respect to x and μ , we obtain

$$\|g - \overline{D}_{n,\delta,\mu}(g)\|_{L_\mu^1} \leq \|g'\|_{C([0,1])} \cdot \|\overline{D}_{n,\delta,\mu}(\varphi_x)\|_{L_\mu^1},$$

which immediately leads to

$$\begin{aligned}
\|f - \overline{D}_{n,\delta,\mu}(f)\|_{L_\mu^1} &\leq 2\|f - g\|_{L_\mu^1} + \|f - g\|_{L_\delta^1} + \|g'\|_{C([0,1])} \cdot \|\overline{D}_{n,\delta,\mu}(\varphi_x)\|_{L_\mu^1} \\
&\leq 2 \left(\|f - g\|_{L_\mu^1} + \|f - g\|_{L_\delta^1} + \|g'\|_{C([0,1])} \cdot \frac{\|\overline{D}_{n,\delta,\mu}(\varphi_x)\|_{L_\mu^1}}{2} \right)
\end{aligned}$$

and to the conclusion of the theorem. \square

In what follows, because of some difference with respect to the case $p = 1$, we extend separately Theorem 3.3 to the L_μ^p space with $p > 1$.

Theorem 3.4. *Let μ be a bounded, monotone, submodular, strictly positive set function on $\mathcal{B}_{[0,1]}$, which also is continuous by increasing sequences of sets, that is if $A_n \in \mathcal{B}_{[0,1]}$, $n \in \mathbb{N}$, with $A_n \subset A_{n+1}$, for all n and $A := \bigcup_{n=1}^\infty A_n \in \mathcal{B}_{[0,1]}$, then $\lim_{n \rightarrow \infty} \mu(A_n) = \mu(A)$.*

Also, let δ be a bounded, strictly positive Borel measure on $\mathcal{B}_{[0,1]}$, such that $\mu(A) \leq \delta(A)$ for all $A \in \mathcal{B}_{[0,1]}$. Then, for any $p > 1$, $L_{\delta,+}^p \subset L_{\mu,+}^p$ and for the Bernstein-Durrmeyer-Choquet operators $\overline{D}_{n,\delta,\mu}(f)(x)$ defined by Theorem 3.3, for all $f \in L_{\delta,+}^p = L_\delta^p \cap \{f : [0, 1] \rightarrow [0, +\infty)\}$ and $n \in \mathbb{N}$, we have

$$\|f - \overline{D}_{n,\delta,\mu}(f)\|_{L_\mu^p} \leq 2\overline{K} \left(f; \frac{\|\overline{D}_{n,\delta,\mu}(\varphi_x)\|_{L_\mu^p}}{2} \right)_{L_{\mu,\delta}^p}.$$

Proof. The proof for $L_{\delta,+}^p \subset L_{\mu,+}^p$ follows exactly as in the proof of Theorem 3.3.

By the convexity of t^p on $[0, +\infty)$, by $\sum_{k=0}^n p_{n,k}(x) = 1$, we easily arrive at the inequalities (exactly as, for example, in the proof of Lemma 2.2 in [13])

$$\begin{aligned} \|\overline{D}_{n,\delta,\mu}(f)\|_{L_\mu^p}^p &\leq (C) \int_0^1 \left[\sum_{k=1}^{n-1} p_{n,k}(x) \cdot \frac{\left(\int_0^1 f(t) p_{n,k}(t) d\delta(t) \right)^p}{\left(\int_0^1 p_{n,k}(t) d\delta(t) \right)^p} \right. \\ &\quad \left. + x^n \cdot \frac{\left((C) \int_0^1 f(t) t^n d\mu(t) \right)^p}{\left((C) \int_0^1 t^n d\mu(t) \right)^p} \right] d\mu(x). \end{aligned}$$

Applying the Hölder's inequality for the integrals from the denominators (in the case of Choquet integrals with respect to μ , the inequality is the same with that for the integrals with respect to the Borel measure δ , see. e.g., Theorem 3.5 in [16] or Theorem 2 in [5]) and reasoning as in the proof of Lemma 2.2 in [13] and as for formula (6) in the proof of Theorem 3.3, we easily arrive at

$$\|\overline{D}_{n,\delta,\mu}(f)\|_{L_\mu^p}^p \leq \|f\|_{L_\mu^p}^p + \|f\|_{L_\delta^p}^p, \text{ for all } f \in L_{\delta,+}^p.$$

Then, since Hölder's inequality for the Choquet integral with respect to μ implies the Minkowski inequality (see, e.g., Theorem 3.7 in [16] or Theorem 2 in [5]), using the above inequality and exactly the reasonings in the proof of Theorem 2.1 in [13], we arrive at the desired inequality in the statement.

It remains to discuss the requirement on μ to be continuous by increasing sequences of sets. This is due to the fact that for Choquet integrals, the Hölder's inequality hold only if both integrals from its right-hand side are not equal to zero (see the proofs of Theorem 3.5 in [16] or of Theorem 2 in [5]).

To have valid the Hölder's inequality in its full generality, we need that for $F \geq 0$, $(C) \int_0^1 F(t) d\mu = 0$ if and only if $F(t) = 0$, μ almost everywhere on $[0, 1]$. But according to Theorem 11.3, p. 228 in [17], if μ is continuous by increasing sequences of sets, then the above mentioned property holds. \square

Remark 3.5. Concrete choices for μ and δ in Theorem 3.3 can be, for example, $\delta(A) = m(A)$ and $\mu(A) = \sin[m(A)]$, where m is the Lebesgue measure on $\mathcal{B}_{[0,1]}$. Indeed, $\mu(A) \leq m(A)$ for all $A \in \mathcal{B}_{[0,1]}$ and since \sin is concave on $[0, 1]$ and $\sin(0) = 0$, by Remark 2.3, (vi) it follows that μ is bounded, monotone and submodular.

Remark 3.6. It is easy to see that another Bernstein-Durrmeyer-Choquet operator satisfying the estimates in Theorems 3.3 and 3.4, can be defined by

$$\tilde{D}_{n,\delta,\mu}(f)(x) = (1-x)^n \frac{(C) \int_0^1 f(t) (1-t)^n d\mu(t)}{(C) \int_0^1 (1-t)^n d\mu(t)} + \sum_{k=1}^n p_{n,k}(x) \frac{\int_0^1 f(t) p_{n,k}(t) d\delta(t)}{\int_0^1 p_{n,k}(t) d\delta(t)}.$$

Also, defining

$$D_{n,\delta,\mu}^*(f)(x) = (1-x)^n \frac{(C) \int_0^1 f(t) (1-t)^n d\mu(t)}{(C) \int_0^1 (1-t)^n d\mu(t)} + \sum_{k=1}^{n-1} p_{n,k}(x) \frac{\int_0^1 f(t) p_{n,k}(t) d\delta(t)}{\int_0^1 p_{n,k}(t) d\delta(t)}$$

$$+x^n \cdot \frac{(C) \int_0^1 f(t)t^n d\mu(t)}{(C) \int_0^1 t^n d\mu(t)},$$

by similar reasonings with those in the proofs of Theorems 3.3 and 3.4, for any $p \geq 1$ we immediately obtain the estimate

$$\|f - D_{n,\delta,\mu}^*(f)\|_{L_\mu^p} \leq 3\overline{K} \left(f; \frac{\|D_{n,\delta,\mu}^*(\varphi_x)\|_{L_\mu^p}}{3} \right).$$

Remark 3.7. For δ a bounded Borel measure on $\mathcal{B}_{[0,1]}$, denote by $D_{n,\delta}$ the classical Bernstein-Durrmeyer operator (i.e. with all the integrals in terms of δ). By Theorem 4.5 in [4], we have the estimate

$$\|D_{n,\delta}(f) - f\|_{L_\delta^1} \leq 2K \left(f; \frac{\|D_{n,\delta}(\varphi_x)\|_{L_\delta^1}}{2} \right)_{L_\delta^1}, \text{ for all } f \in L_\delta^1,$$

where $K(f; t)_{L_\delta^1} = \inf_{g \in C^1([0,1])} \{\|f - g\|_{L_\delta^1} + t\|g'\|_{C([0,1])}\}$.

Comparing with the estimate for $\|\overline{D}_{n,\delta,\mu}(f) - f\|_{L_\mu^1}$ in Theorem 3.3 and taking into account that $\|f - g\|_{L_\mu^1} \leq \|f - g\|_{L_\delta^1}$ and $\|D_{n,\delta}(\varphi_x)\|_{L_\mu^1} \leq \|D_{n,\delta}(\varphi_x)\|_{L_\delta^1}$, it follows that it is possible that in some cases, $\overline{D}_{n,\delta,\mu}(f)$ in Theorem 3.3, approximates better $f \in L_{\delta,+}^1$ in the L_μ^1 -norm than approximate $D_{n,\delta}(f)$ the same function f but in the L_δ^1 -norm.

4 Concrete Bernstein-Durrmeyer-Choquet operators

Since the estimates in Theorem 3.1 are of very general and abstract form, involving the apparently difficult to be calculated Choquet integrals, it is of interest to obtain in some particular cases, concrete expressions for the order of approximation.

In this sense, we will apply Theorem 3.1, (i), for $d = 1$ and for some special choices of the submodular set functions.

Thus, we will consider the case of the measures of possibility. Denoting $p_{n,k}(x) = \binom{n}{k} x^k (1-x)^{n-k}$, let us define $\lambda_{n,k}(t) = \frac{p_{n,k}(t)}{k^k n^{-n} (n-k)^{n-k} \binom{n}{k}} = \frac{t^k (1-t)^{n-k}}{k^k n^{-n} (n-k)^{n-k}}$, $k = 0, \dots, n$. Here, by convention we consider $0^0 = 1$, so that the cases $k = 0$ and $k = n$ have sense.

By considering the root $\frac{k}{n}$ of $p'_{n,k}(x)$, it is easy to see that $\max\{p_{n,k}(t); t \in [0, 1]\} = k^k n^{-n} (n-k)^{n-k} \binom{n}{k}$, which implies that each $\lambda_{n,k}$ is a possibility distribution on $[0, 1]$. Denoting by $P_{\lambda_{n,k}}$ the possibility measure induced by $\lambda_{n,k}$ and $\Gamma_{n,x} := \Gamma_n = \{P_{\lambda_{n,k}}\}_{k=0}^n$ (i.e. Γ is independent of x), the nonlinear Bernstein-Durrmeyer polynomial operators given by (2), in terms of the Choquet integrals with respect to the set functions in Γ_n , will become

$$D_{n,\Gamma_n}(f)(x) = \sum_{k=0}^n p_{n,k}(x) \cdot \frac{(C) \int_0^1 f(t) t^k (1-t)^{n-k} dP_{\lambda_{n,k}}(t)}{(C) \int_0^1 t^k (1-t)^{n-k} dP_{\lambda_{n,k}}(t)}. \quad (7)$$

It is easy to see that any possibility measure $P_{\lambda_{n,k}}$ is bounded, monotone, sub-modular and strictly positive, $n \in \mathbb{N}$, $k = 0, 1, \dots, n$, so that we are under the hypothesis of Theorem 3.1, (i).

We can state the following result.

Theorem 4.1. *If $D_{n,\Gamma_n}(f)(x)$ is given by (7), then for every $f \in C_+([0, 1])$, $x \in [0, 1]$ and $n \in \mathbb{N}$, $n \geq 2$, we have*

$$|D_{n,\Gamma_n}(f)(x) - f(x)| \leq 2\omega_1 \left(f; \frac{(1 + \sqrt{2})\sqrt{x(1-x)} + \sqrt{2}\sqrt{x}}{\sqrt{n}} + \frac{1}{n} \right)_{[0,1]}.$$

For its proof, we need the following auxiliary result.

Lemma 4.2. *Let $n \in \mathbb{N}$, $n \geq 2$ and $x \in [0, 1]$. Denoting*

$$A_{n,k}(x) := \sup\{|t - x|t^k(1 - t)^{n-k}; t \in [0, 1]\} =$$

$$\max\{\sup\{(t - x)t^k(1 - t)^{n-k}; t \in [x, 1]\}, \sup\{(x - t)t^k(1 - t)^{n-k}; t \in [0, x]\}\},$$

with the convention $0^0 = 1$, for all $k = 0, \dots, n$ we have

$$A_{n,k}(x) = \max\{(t_2 - x)t_2^k(1 - t_2)^{n-k}, (x - t_1)t_1^k(1 - t_1)^{n-k}\},$$

with t_1, t_2 given by

$$t_1 = \frac{nx + k + 1 - \sqrt{\Delta}}{2(n + 1)}, \quad t_2 = \frac{nx + k + 1 + \sqrt{\Delta}}{2(n + 1)}, \quad (8)$$

where

$$\begin{aligned} \Delta &= (nx + k + 1)^2 - 4kx(n + 1) = n^2 \left[(x + (k + 1)/n)^2 - 4x \frac{k}{n} \cdot \frac{n + 1}{n} \right] \\ &= (nx - k)^2 + 2x(n - k) + 2k(1 - x) + 1 \geq 1. \end{aligned}$$

Proof. Let us denote $H_{n,k}(t) = t^k(1 - t)^{n-k}|t - x|$, with $k \in \{0, \dots, n\}$. We have two cases : (i) $1 \leq k \leq n - 1$ and (ii) $k = 0$ or $k = n$.

Case (i). For $t \in [x, 1]$ we obtain $H_{n,k}(t) = (t - x)t^k(1 - t)^{n-k}$ and from $H'_{n,k}(t) = t^{k-1}(1 - t)^{n-k-1}[-t^2(n + 1) + t(nx + k + 1) - kx] = 0$, it follows $-t^2(n + 1) + t(nx + k + 1) - kx = 0$, which has the discriminant

$$\Delta = (nx + k + 1)^2 - 4kx(n + 1) = (nx - k)^2 + 2x(n - k) + 2k(1 - x) + 1 \geq 1.$$

Therefore, the quadratic equation has two real distinct solutions $t_1 < t_2$

$$t_1 = \frac{nx + k + 1 - \sqrt{\Delta}}{2(n + 1)}, \quad t_2 = \frac{nx + k + 1 + \sqrt{\Delta}}{2(n + 1)},$$

where by simple calculation we derive $0 \leq t_1 < t_2 \leq 1$. Also, since $H_{n,k}(0) = H_{n,k}(x) = H_{n,k}(1) = 0$ and $H_{n,k}(t) \geq 0$ for $t \in [x, 1]$, simple graphical reasonings show that the only possibility is $0 \leq t_1 \leq x \leq t_2 \leq 1$, with $t = t_2$ maximum point on $[x, 1]$ for $H_{n,k}(t)$.

Similarly, for $t \in [0, x]$, since $H_{n,k}(t) = (x-t)t^k(1-t)^{n-k}$, using the above reasonings we obtain $H'_{n,k}(t) = t^{k-1}(1-t)^{n-k-1}[t^2(n+1) - t(nx+k+1) + kx]$ and that t_1 is a maximum point of $H_{n,k}(t)$ on $[0, x]$.

In conclusion, with t_1, t_2 given by (8), we get

$$A_{n,k}(x) = \max\{(t_2 - x)t_2^k(1 - t_2)^{n-k}, (x - t_1)t_1^k(1 - t_1)^{n-k}\}.$$

Case (ii). Suppose first that $k = 0$. By the calculation from the case (i), for $t \in [x, 1]$ we get $0 = t_1 \leq x \leq t_2 = \frac{nx+1}{n+1} \leq 1$, $H_{n,0}(t) \geq 0$ and $H_{n,0}(x) = H_{n,0}(1) = 0$, which by similar graphical reasonings leads to the fact that the maximum of $H_{n,0}(t)$ on $[x, 1]$ is $H_{n,0}(t_2) = (t_2 - x)(1 - t_2)^n$. Therefore, we recapture the case (i) with the convention that $0^0 = 1$. Similarly, for $t \in [0, x]$, we get that the maximum of $H_{n,0}(t)$ is $H_{n,0}(t_1) = (x - t_1)(1 - t_1)^n$.

The subcase $k = n$ is similar, which proves the lemma. \square

Proof of Theorem 4.1. According to Theorem 3.1, (i), we have to estimate

$$D_{n,\Gamma_n}(\varphi_x)(x) = \sum_{k=0}^n p_{n,k}(x) \cdot \frac{(C) \int_0^1 |t-x| t^k (1-t)^{n-k} dP_{\lambda_{n,k}}(t)}{(C) \int_0^1 t^k (1-t)^{n-k} dP_{\lambda_{n,k}}(t)}.$$

First of all, by Definition 2.2, (ii), we get

$$\begin{aligned} (C) \int_0^1 t^k (1-t)^{n-k} dP_{\lambda_{n,k}}(t) &= \int_0^{+\infty} P_{\lambda_{n,k}}(\{t \in [0, 1]; t^k(1-t)^{n-k} \geq \beta\}) d\beta \\ &= \int_0^1 P_{\lambda_{n,k}}(\{t \in [0, 1]; t^k(1-t)^{n-k} \geq \beta\}) d\beta \\ &= \int_0^1 \sup\{\lambda_{n,k}(s); s \in \{t \in [0, 1]; t^k(1-t)^{n-k} \geq \beta\}\} d\beta \\ &= \frac{1}{k^k n^{-n} (n-k)^{n-k}} \cdot \int_0^1 \sup\{s^k(1-s)^{n-k}; s \in \{t \in [0, 1]; t^k(1-t)^{n-k} \geq \beta\}\} d\beta. \end{aligned}$$

For simplicity, denote $E_{n,k} = k^k n^{-n} (n-k)^{n-k}$, where again we take $0^0 = 1$. Since for $\beta > E_{n,k}$ we have $\{t \in [0, 1]; t^k(1-t)^{n-k} \geq \beta\} = \emptyset$ and since we can take $\sup\{s^k(1-s)^{n-k}; s \in \emptyset\} = 0$, it follows

$$\begin{aligned} (C) \int_0^1 t^k (1-t)^{n-k} dP_{\lambda_{n,k}}(t) &= \frac{1}{E_{n,k}} \cdot \int_0^{E_{n,k}} \sup\{s^k(1-s)^{n-k}; s \in \{t \in [0, 1]; t^k(1-t)^{n-k} \geq \beta\}\} d\beta \\ &= \frac{1}{E_{n,k}} \cdot \int_0^{E_{n,k}} E_{n,k} d\beta = E_{n,k}. \end{aligned} \tag{9}$$

On the other hand, denoting $A_{n,k}(x) = \sup\{|t-x|t^k(1-t)^{n-k}; t \in [0,1]\}$, by Remark 2.3, (iii), (v) and by Lemma 4.2, for $t_1 < t_2$ in (8) we obtain

$$\begin{aligned} (C) \int_0^1 |t-x|t^k(1-t)^{n-k} dP_{\lambda_{n,k}}(t) &\leq (C) \int_0^1 A_{n,k}(x) dP_{\lambda_{n,k}}(t) \\ &= A_{n,k}(x)(C) \int_0^1 1 dP_{\lambda_{n,k}}(t) = \max\{(t_2-x)t_2^k(1-t_2)^{n-k}, (x-t_1)t_1^k(1-t_1)^{n-k}\} \\ &\leq (t_2-x)t_2^k(1-t_2)^{n-k} + (x-t_1)t_1^k(1-t_1)^{n-k}. \end{aligned}$$

Since $\frac{t_2^k(1-t_2)^{n-k}}{k^k n^{-n}(n-k)^{n-k}} \leq 1$, $\frac{t_1^k(1-t_1)^{n-k}}{k^k n^{-n}(n-k)^{n-k}} \leq 1$ and by Lemma 4.2 we get

$$\begin{aligned} \frac{A_{n,k}(x)}{k^k n^{-n}(n-k)^{n-k}} &\leq (t_2-x) \cdot \frac{t_2^k(1-t_2)^{n-k}}{k^k n^{-n}(n-k)^{n-k}} + (x-t_1) \cdot \frac{t_1^k(1-t_1)^{n-k}}{k^k n^{-n}(n-k)^{n-k}} \\ &\leq t_2 - t_1 = \frac{\sqrt{\Delta}}{n+1} \leq \frac{\sqrt{(nx-k)^2 + 2x(n-k) + 2k(1-x) + 1}}{n} \\ &\leq \sqrt{(x-k/n)^2 + 2x/n + (2k/n) \cdot (1-x)/n + 1/n^2} \\ &\leq |x-k/n| + \sqrt{2x}/\sqrt{n} + (\sqrt{2k}/\sqrt{n}) \cdot \sqrt{(1-x)/n} + 1/n, \end{aligned}$$

this immediately implies

$$\begin{aligned} D_{n,\Gamma_n}(\varphi_x)(x) &\leq \sum_{k=0}^n p_{n,k}(x)(|x-k/n| + \sqrt{2x}/\sqrt{n} + \sqrt{2k/n} \cdot \sqrt{(1-x)/n} + 1/n) \\ &\leq \frac{\sqrt{x(1-x)}}{\sqrt{n}} + \frac{\sqrt{2x}}{\sqrt{n}} + \frac{\sqrt{2}\sqrt{x(1-x)}}{\sqrt{n}} + \frac{1}{n} = \frac{(1+\sqrt{2})\sqrt{x(1-x)} + \sqrt{2}\sqrt{x}}{\sqrt{n}} + \frac{1}{n}. \end{aligned}$$

Above we have used the well-known estimate $\sum_{k=0}^n p_{n,k}(x)|x-k/n| \leq \frac{\sqrt{x(1-x)}}{\sqrt{n}}$ and the Cauchy-Schwarz inequality for Bernstein polynomials, $B_n(f)(x) \leq \sqrt{B_n(f^2)(x)}$, applied for $f(t) = \sqrt{t}$.

Finally, applying Theorem 3.1, (i), the proof of Theorem 4.1 follows. \square

Remark 4.3. For $\mu = \sqrt{m}$ with m denoting the Lebesgue measure on $[0,1]$, another particular case of the Bernstein-Durrmeyer operators for which the quantitative estimates in Theorem 3.1 are applicable would be, for example,

$$D_{n,\mu}(f)(x) = \sum_{k=0}^n p_{n,k}(x) \cdot \frac{(C) \int_0^1 f(t)t^k(1-t)^{n-k} d\mu}{(C) \int_0^1 t^k(1-t)^{n-k} d\mu}.$$

It is worth noting that the uniform convergence of this $D_{n,\mu}(f)$ to f , follows directly from the general Theorem 3.2 in [10]. Also, it can be obtained by using the nonlinear Feller kind scheme expressed by Theorem 3.1 in [9] (combined with Remark 3.2 there), since by direct calculation we can show that $D_{n,\mu}(e_1)$ converges uniformly to e_1 and $D_{n,\mu}((\cdot-x)^2)$ converges uniformly to 0, on $[0,1]$.

Remark 4.4. Since the Bernstein-Durrmeyer-Choquet operators in this paper can be defined with respect to a family of Borel or Choquet measures, combined in various ways, this fact offers a very high flexibility and generality, allowing to construct operators having even better approximation properties. A first example for this flexibility is shown by Theorem 3.4 and Remark 3.6.

For the second example, let us replace in formula (7) the family Γ_n of measures of possibilities $P_{\lambda_{n,k}}$, $k = 0, \dots, n$, by the family consisting in the Dirac measures $\delta_{k/n}$, $k = 0, 1, \dots, n-1$, (which are Borel measures and therefore with the corresponding Choquet integrals reducing to the classical ones) together with a monotone, submodular, strictly positive set function μ . Then, denoting by $B_n(f)(x)$ the classical Bernstein operators, for D_{n,Γ_n} in (7) we get

$$\begin{aligned} D_{n,\Gamma_n}(f)(x) - f(x) &= \left[\sum_{k=0}^{n-1} p_{n,k}(x) f\left(\frac{k}{n}\right) + x^n \cdot \frac{(C) \int_0^1 f(t) t^n d\mu(t)}{(C) \int_0^1 t^n d\mu(t)} \right] - f(x) \\ &= B_n(f)(x) - f(x) + x^n \left[\frac{(C) \int_0^1 f(t) t^n d\mu(t)}{(C) \int_0^1 t^n d\mu(t)} - f(1) \right]. \end{aligned}$$

Suppose now that $f \geq 0$ is strictly increasing and strictly convex on $[0, 1]$ and, for example, that $\mu(A) = \sqrt{m(A)}$ or $\mu(A) = \sin[m(A)]$, with m the Lebesgue measure. The strict convexity implies $B_n(f)(x) - f(x) > 0$ for all $x \in (0, 1)$ and the property of f to be strictly increasing easily implies

$$\frac{(C) \int_0^1 f(t) t^n d\mu(t)}{(C) \int_0^1 t^n d\mu(t)} - f(1) < \frac{f(1) \cdot (C) \int_0^1 t^n d\mu(t)}{(C) \int_0^1 t^n d\mu(t)} - f(1) = 0.$$

Therefore, in this case we get

$$|D_{n,\Gamma_n}(f)(x) - f(x)| < \max \left\{ |B_n(f)(x) - f(x)|, x^n \left| \frac{(C) \int_0^1 f(t) t^n d\mu(t)}{(C) \int_0^1 t^n d\mu(t)} - f(1) \right| \right\},$$

i.e. for $x \in (0, 1)$, $D_{n,\Gamma_n}(f)(x)$ approximates better than $B_n(f)(x)$.

Here it is clear that $B_n(f)(x)$ can also be viewed as the Bernstein-Durrmeyer operators in the case when Γ_n is composed by the Dirac measures $\delta_{k/n}$, $k = 0, \dots, n$, where we note that although the Dirac measures are not strictly positive, however the Bernstein-Durrmeyer operators attached to them are well defined. This fact contrasts with the classical case in [1] when Γ_n is composed by only one set function, independent of n , and when the strict positivity of the set function is necessarily for the convergence (see Theorem 1 in [1]). In other words, the strict positivity of the set functions in Theorem 3.1 is not always necessary.

For another example, let us consider the genuine Bernstein-Durrmeyer-Choquet operators given by

$$U_{n,\Gamma_n}(f)(x) = p_{n,0}(x) \cdot \frac{(C) \int_0^1 f(t)(1-t)^n d\nu_{n,0}}{(C) \int_0^1 (1-t)^n d\nu_{n,0}} + p_{n,n}(x) \cdot \frac{(C) \int_0^1 f(t)t^n d\nu_{n,n}}{(C) \int_0^1 t^n d\nu_{n,n}}$$

$$+ \sum_{k=1}^{n-1} p_{n,k}(x) \cdot \frac{(C) \int_0^1 f(t) p_{n-2,k-1}(t) d\mu_{n-2,k-1}(t)}{(C) \int_0^1 p_{n-2,k-1}(t) d\mu_{n-2,k-1}(t)},$$

where $\Gamma_n = \{\nu_{n,0}, \nu_{n,n}, \mu_{n-2,k-1}, k = 1, \dots, n-1\}$.

Let us denote by $G_n(f)(x)$, the classical genuine Bernstein-Durrmeyer operator (see, e.g., [11]). Choosing in Γ_n the set functions $\mu_{n-2,k-1}, k = 1, \dots, n-1$ as the Lebesgue measure, $\nu_{n,0} = \delta_0$ (as Dirac measure) and $\nu_{n,n}$ as a monotone, submodular and strictly positive set function, we immediately obtain

$$U_{n,\Gamma_n}(f)(x) - f(x) = G_n(f)(x) - f(x) + x^n \left[\frac{(C) \int_0^1 f(t) t^n d\nu_{n,n}(t)}{(C) \int_0^1 t^n d\nu_{n,n}(t)} - f(1) \right].$$

Since the strict convexity of f implies $G_n(f)(x) - f(x) > 0$ for all $x \in (0, 1)$ (see, e.g., Lemma 2.1, (iv) in [11]), similar reasonings with those for the previous example show that if f is strictly convex and strictly increasing on $[0, 1]$ (and, for example, $\nu_{n,n}(A) = \sqrt{m(A)}$ or $\nu_{n,n}(A) = \sin[m(A)]$), then $U_{n,\Gamma_n}(f)(x)$ approximates better f on $(0, 1)$ than the classical genuine operator, $G_n(f)(x)$.

Remark 4.5. Recall that in [9], Example 4.2, for the nonlinear Picard-Choquet operators we have obtained a general estimate similar to that for the classical Picard operators, while for particular functions of the form $f(x) = Me^{-Ax}$, $M, A > 0$, we got there essentially better error estimates.

Remark 4.6. In [13] applications of the classical Bernstein-Durrmeyer operators to learning theory are presented. Taking onto account the very recent applications of the Choquet integral to learning theory (see, e.g., [14] and the references therein), it becomes of interest to see for potential applications of the Bernstein-Durrmeyer-Choquet operators to learning theory. Also, taking into account the applications of the classical Bernstein-Durrmeyer operators in regression estimation in, e.g., [15] and the very recent applications of the Choquet integral to regression model, see, e.g., [12], it would be of interest to see for possible applications of the Bernstein-Durrmeyer-Choquet operators to the regression model.

References

- [1] Berdysheva, E.E. : Uniform convergence of Bernstein-Durrmeyer operators with respect to arbitrary measure. J. Math. Anal. Appl. **394**, 324-336(2012)
- [2] Berdysheva, E.E. : Bernstein-Durrmeyer operators with respect to arbitrary measure. II : Pointwise convergence. J. Math. Anal. Appl. **418**, 734-752(2014)
- [3] Berdysheva, E.E., Li, B.-Z. : On L^p -convergence of Bernstein-Durrmeyer operators with respect to arbitrary measure. Publ. Inst. Math. (Beograd). N.S. **96**(110), 23-29(2014)
- [4] Berdysheva, E.E., Jetter, K. : Multivariate Bernstein-Durrmeyer operators with arbitrary weight functions. J. Approx. Theory **162**, 576-598(2010)

- [5] Cerdà, J., Martín, J., Silvestre, P. : Capacitary function spaces, *Collect. Math.*, **62**, 95-118(2011)
- [6] Choquet, G. : Theory of capacities. *Ann. Inst. Fourier (Grenoble)* **5**, 131-295(1954)
- [7] Denneberg, D. : *Non-Additive Measure and Integral*. Kluwer Academic Publisher. Dordrecht(1994)
- [8] Dubois, D., Prade, H. : *Possibility Theory*. Plenum Press. New York(1988)
- [9] Gal, S.G. : Approximation by Choquet integral operators. *Ann. Mat. Pura Appl.* online access. DOI : 10.1007/s10231-015-0495-x
- [10] Gal, S.G., Opris, B.D. : Uniform and pointwise convergence of Bernstein-Durrmeyer operators with respect to monotone and submodular set functions. *J. Math. Anal. Appl.* **424**, 1374-1379(2015)
- [11] Gonska, H., Kacsó D., Rasa, I. : The genuine Bernstein-Durrmeyer operators revisited. *Results Math.* **62**, 295-310(2012)
- [12] Grabisch, M. : Modelling data by the Choquet integral. In : *Information Fusion in Data Mining* (Torra V., ed.). Physica Verlag. Heidelberg, pp. 135-148(2003)
- [13] Li, B.-Z. : Approximation by multivariate Bernstein-Durrmeyer operators and learning rates of least-square regularized regression with multivariate polynomial kernel. *J. Approx. Theory* **173**, 33-55(2013)
- [14] Tehrani, A.F., Cheng, W., Dembczyński, K., Hüllermeier, E. : Learning monotone nonlinear models using Choquet integrals. *Machine Learning*. **89**, 183-211(2012)
- [15] Rafajlowicz, E., Skubalska-Rafajlowicz, E. : Nonparametric regression estimation by Bernstein-Durrmeyer polynomials. *Tatra Mt. Mat. Publ.* **17**, 227-239(1999)
- [16] Wang, R.S. : Some inequalities and convergence theorems for Choquet integrals, *J. Appl. Math. Comput.* **35**, 305-321(2011)
- [17] Wang, Z., Klir, G.J. : *Generalized Measure Theory*. Springer. New York(2009)